

# Analysis of the Size Distributions of Fetal and Maternal Cell-Free DNA by Paired-End Sequencing

H. Christina Fan,<sup>1</sup> Yair J. Blumenfeld,<sup>2</sup> Usha Chitkara,<sup>2</sup> Louanne Hudgins,<sup>3</sup> and Stephen R. Quake<sup>1\*</sup>

**BACKGROUND:** Noninvasive prenatal diagnosis with cell-free DNA in maternal plasma is challenging because only a small portion of the DNA sample is derived from the fetus. A few previous studies provided size-range estimates of maternal and fetal DNA, but direct measurement of the size distributions is difficult because of the small quantity of cell-free DNA.

**METHODS:** We used high-throughput paired-end sequencing to directly measure the size distributions of maternal and fetal DNA in cell-free maternal plasma collected from 3 typical diploid and 4 aneuploid male pregnancies. As a control, restriction fragments of  $\lambda$  DNA were also sequenced.

**RESULTS:** Cell-free DNA had a dominant peak at approximately 162 bp and a minor peak at approximately 340 bp. Chromosome Y sequences were rarely longer than 250 bp but were present in sizes of <150 bp at a larger proportion compared with the rest of the sequences. Selective analysis of the shortest fragments generally increased the fetal DNA fraction but did not necessarily increase the sensitivity of aneuploidy detection, owing to the reduction in the number of DNA molecules being counted. Restriction fragments of  $\lambda$  DNA with sizes between 60 bp and 120 bp were preferentially sequenced, indicating that the shotgun sequencing work flow introduced a bias toward shorter fragments.

**CONCLUSIONS:** Our results confirm that fetal DNA is shorter than maternal DNA. The enrichment of fetal DNA by size selection, however, may not provide a dramatic increase in sensitivity for assays that rely on length measurement in situ because of a trade-off between the fetal DNA fraction and the number of molecules being counted.

© 2010 American Association for Clinical Chemistry

Traditional methods of prenatal diagnosis of genetic disorders use materials obtained by amniocentesis or chorionic villus sampling, invasive procedures that carry a small but clear risk of miscarriage (1). The discovery of cell-free fetal nucleic acids in the plasma of pregnant mothers has led to the development of several noninvasive prenatal diagnostic techniques in the past decade (2). The detection of fetal aneuploidy and autosomal recessive disorders with cell-free nucleic acids is particularly challenging because only a small portion of the cell-free nucleic acids in maternal plasma is derived from the fetus. We recently demonstrated noninvasive detection of fetal aneuploidy by high-throughput shotgun sequencing of cell-free DNA (3), and an independent group quickly reproduced our results (4, 5). For almost all prenatal diagnostic assays, the background of maternal DNA provides a practical limit on sensitivity, and therefore the fraction of fetal DNA present in the maternal plasma is a critical parameter. There is evidence that fetal DNA is shorter on balance than maternal DNA, and therefore substantial effort has been invested in developing methods to enrich for fetal DNA (6, 7). Extracting fractions of lower molecular weight DNA with electrophoretic techniques or the use of smaller PCR amplicons could increase the fraction of fetal DNA, and such methods have been used to improve the detection of fetal point mutations and the determination of fetal genotypes (8–12).

Paired-end sequencing is a technique that obtains sequence information for both ends of each DNA molecule. By finding the coordinates of the 2 sequences on the genome through sequence alignment, one can deduce the length of the DNA fragment. A single sequencing experiment yields sequence and size information for tens of millions of DNA fragments. In this study, we used high-throughput paired-end sequencing of cell-free DNA in maternal plasma to study the length distributions of fetal and maternal DNA. Paired-end sequencing enabled us to directly measure

<sup>1</sup> Department of Bioengineering, Stanford University and Howard Hughes Medical Institute, Stanford, CA; <sup>2</sup> Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Stanford University, Stanford, CA; <sup>3</sup> Division of Medical Genetics, Department of Pediatrics, Stanford University, Stanford, CA.  
\* Address correspondence to this author at: Department of Bioengineering,

Stanford University and Howard Hughes Medical Institute, 318 Campus Dr., Clark Center, Room E300, Stanford, CA 94305. Fax 650-7361961; e-mail quake@stanford.edu.

Received January 28, 2010; accepted May 10, 2010.

Previously published online at DOI: 10.1373/clinchem.2010.144188

---

the size distributions of maternal DNA and fetal DNA with single-base resolution from cell-free DNA collected from women carrying male fetuses, without the need to pool samples and with much higher precision than can be obtained by gel electrophoresis or via the PCR. Our data confirm that fetal DNA is shorter than maternal DNA and is predominantly within the size range of a mononucleosome. We demonstrated that the shotgun sequencing work flow introduces a bias toward shorter fragments, a phenomenon that effectively enriches the fetal DNA fraction. Finally, by selectively analyzing only the shortest fragments, we showed that there is a delicate trade-off in sensitivity in fetal aneuploidy detection between the fetal DNA fraction and the number of molecules counted.

## Materials and Methods

### SAMPLE PROCESSING

Blood samples were collected at the Lucile Packard Children's Hospital (Stanford University), with informed consent obtained under an institutional review board-approved study. Maternal blood samples from 7 pregnancies with male fetuses, including 2 cases of trisomy 21, a case of trisomy 13, and a case of trisomy 18, were selected for this study. These samples were collected at gestational ages of 12–23 weeks. Plasma was first separated from the blood cells by centrifugation at 1600g at 4 °C for 10 min. The plasma was then centrifuged at 16 000g for 10 min at room temperature to remove residual cells. DNA was extracted from 1.6–2.4 mL of cell-free plasma with the NucleoSpin Plasma F Kit (Macherey-Nagel). To measure the quantity of cell-free DNA, we performed real-time TaqMan PCR assays specific for a chromosome 1 locus and a chromosome Y locus.

To investigate the fragment length-dependent sequencing bias, we prepared a restriction digest of  $\lambda$  DNA (Invitrogen).  $\lambda$  DNA was digested with *AluI*, a 4-bp cutter, for 2 h at 37 °C. The digest was then heated at 65 °C for 20 min to inactivate the enzyme. The digest was purified with the aid of a QIAquick PCR Purification Kit (Qiagen), and 5 ng of the purified DNA was used to construct the sequencing library.

### SEQUENCING LIBRARY CONSTRUCTION

A combination of the protocols detailed in Kozarewa et al. (13) and Fan et al. (3) were used to construct Solexa sequencing libraries. To preserve the original length of plasma DNA, we performed no fragmentation procedures. Full-length paired-end sequencing adaptors were ligated directly onto end-polished, A-tailed double-stranded plasma DNA. The adaptors were purified by HPLC and treated with  $T_4$  polynucleotide kinase to phosphorylate the 5' ends. The final concentra-

tion of the adaptors in the ligation reaction was 800 pmol/L. The libraries were amplified with 12 cycles of the PCR. No agarose gel purification was performed. A Bioanalyzer (Agilent Technologies) and the High Sensitivity DNA Assay were used to analyze the libraries. The libraries were quantified by traditional real-time TaqMan PCR assays with human-specific primers and by digital PCR (Fluidigm) with a universal template assay (14) designed for paired-end libraries. Details of the library-preparation protocols and adaptor sequences can be found in the Data Supplement files that accompany the online version of this article at <http://www.clinchem.org/content/vol56/issue8>.

### SEQUENCING

Libraries were sequenced on the Genome Analyzer II (Illumina) according to the manufacturer's instructions. Thirty-two bases at each end were sequenced.

### SEQUENCE ALIGNMENT

Image analysis, base calling, and alignment were performed with Illumina's Pipeline software (version 1.4.0). For the plasma DNA libraries, we used the ELAND\_PAIR option to map the first 25 bases of each sequenced end to the reference human genome (NCBI Build 36).

For the  $\lambda$  DNA digest library, we performed image analysis, base calling, and alignment for cycles 3–27 and cycles 39–64. The first 2 cycles on both ends were omitted because they corresponded to the restriction site sequences and because the domination of certain bases in the first cycle caused calibration problems in the image analysis software. The sequences were mapped to the genome of  $\lambda$  DNA (GenBank accession no. J02459).

The Pipeline software extracted the data from the *s\_N\_R\_export.txt* files output. These files provided information that included the sequence of a read, the chromosome, the coordinate on the forward strand to which the 5' end of a read mapped with at most 2 mismatches, and the coordinate offset if the paired read also mapped to the same chromosome.

### DATA ANALYSIS

Custom Python and MATLAB scripts were written for further analysis of the data. The absolute value of the coordinate offset plus 25 bases was interpreted as the length of the sequenced DNA fragment. We used only reads that had one end mapped to the forward strand and one end mapped to the reverse strand. In addition, for paired reads with the first read mapped to the forward strand, the offset value in principle should be  $>0$ , whereas for paired reads with the first read mapped to the reverse strand, the offset value should be  $<0$  (see

Fig. 1 in the online Data Supplement). Reads that did not follow this rule were filtered out.

For  $\lambda$  DNA sequences, we counted the number of reads mapped to each restriction site and ignored sites with restriction fragment lengths of  $<30$  bp (because 25 bp was used for alignment). The data were divided into 20-bp bins from 30 bp to 2500 bp. For each 20-bp bin, we calculated the number of reads for all restriction digest fragments falling within the 20-bp bin and divided it by the number of restriction digests within the bin. We fitted the data by locally weighted logistic regression.

To measure the length distributions of maternal and fetal DNA, we tallied the number of reads that had sizes between 30 bp and 510 bp in 20-bp intervals for each chromosome. We applied weighting to each data point by using the fitted data of  $\lambda$  DNA to correct for the length-dependent sequencing bias. For each 20-bp bin, we calculated the -fold increase in fetal DNA fraction as:

$$\frac{f_i / \sum_i f_i}{t_i / \sum_i t_i},$$

where  $f_i$  is the count of fetal (chromosome Y) sequences within the  $i$ th bin and  $t_i$  is the count of all sequences within the  $i$ th bin.

As in our previous study (3), we observed a GC bias in read coverage. To reduce the effect of such bias, we followed the procedures outlined by Fan and Quake (15). Overrepresentation and underrepresentation of chromosomes were measured, and the fetal fraction was estimated from the depletion of chromosome X sequences and/or the overabundance of chromosome 18, 13, or 21, as described in our previous study (3).

## Results

### ANALYSIS OF LENGTH-DEPENDENT BIAS OF ILLUMINA/SOLEXA SEQUENCING

We used the restriction digest of  $\lambda$  DNA to study the effects of library preparation and sequencing on the length distribution of DNA. We prepared a sequencing library from *Alu*I-digested  $\lambda$  DNA that had a total DNA amount similar to that of the plasma DNA samples. Sequencing on a single lane of the flow cell yielded approximately  $14 \times 10^6$  paired-end reads, 97% of which were mapped to restriction sites with the predicted fragment length and used for subsequent analysis (see Table 1 in the online Data Supplement). Fig. 1 is a plot of the number of reads vs restriction fragment length. Bins with 60–120 bp had the most reads. The number

of reads decreased rapidly as the fragment size increased. Very few fragments  $>1$  kb were sequenced.

### SIZE DISTRIBUTION OF TOTAL AND FETAL DNA IN MATERNAL PLASMA DETERMINED BY PAIRED-END SEQUENCING

With real-time PCR, we determined the concentrations of cell-free plasma DNA in the 7 sequenced samples to be within 0.7–5.6  $\mu\text{g/L}$  plasma (assuming 6.6 pg/genome). DYS14, a chromosome Y-specific sequence, was detected in all samples from male fetus pregnancies and was not detected in a female genomic DNA control.

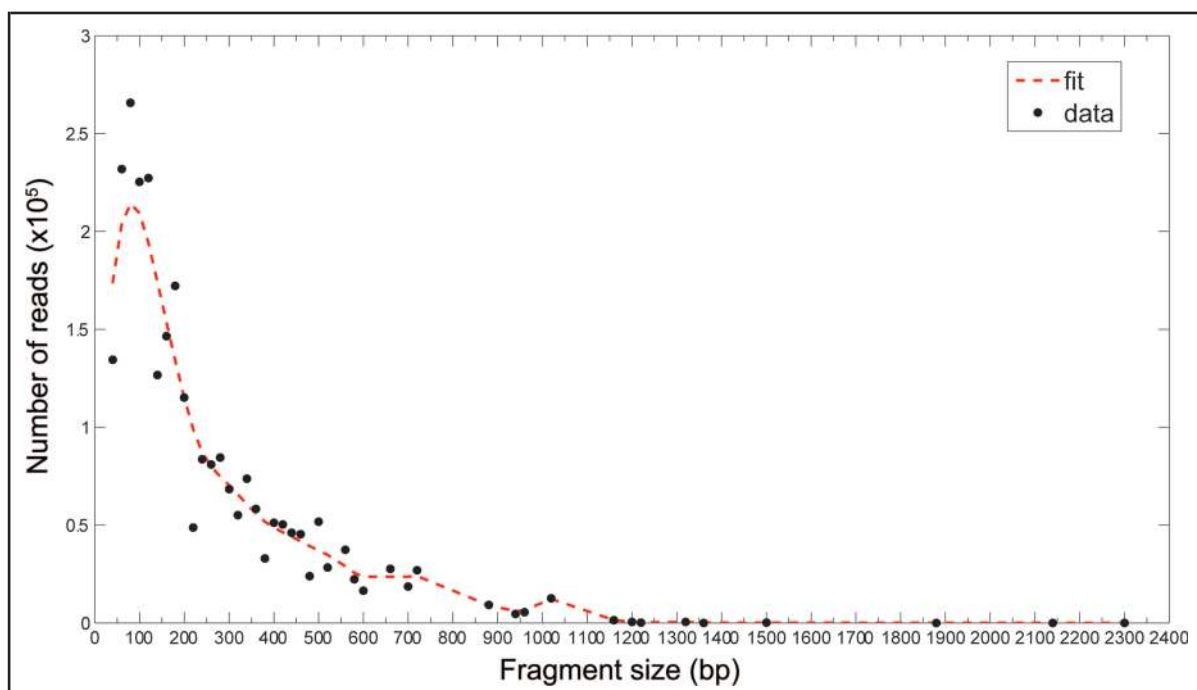
Table 2 in the online Data Supplement presents statistics for the paired-end sequencing run and details of the plasma samples. The mean number of total reads was approximately  $19 \times 10^6$ , with about 52% (i.e.,  $10 \times 10^6$  reads) having both ends mapped to 2 unique locations on a single chromosome with no more than 2 mismatches. Paired-end reads mapped to the forward and reverse strands in equal proportions. We filtered out reads that had ends mapped to the same strand and reads that did not have reasonable offset values (i.e., values that were too large compared with the upper limit of the amplicon size for a PCR reaction). The remaining reads (approximately 99.5% of all paired reads) were used for downstream analyses.

The mean number of chromosome Y reads was approximately 13 000, which is equivalent to approximately 0.1% of the total paired-end reads. Fig. 2 presents the size distribution of sequenced cell-free DNA according to the chromosomes. Sizes ranged from 30 to 510 bp in 20-bp bins. The median length was 162 bp. We applied weighting to the length distribution by using values of the Loess fit from Fig. 1. The dominant peak was approximately 162 bp, approximately the size of a monochromatosome. A minor peak at approximately 340 bp, approximately the size of a dichromatosome, was also observed.

We observed that the size distribution for chromosome Y was shifted for most samples toward the shorter end, compared with the other chromosomes (Fig. 2). Very few chromosome Y sequences had the length of a dichromatosome. Additionally, there were slightly more chromosome Y sequences with lengths shorter than that of a monochromatosome. One can enrich the fraction of fetal DNA by a factor of approximately 1.5 by targeting sequences shorter than 150 bp (Fig. 3).

### FETAL DNA FRACTION AND ANEUPLOIDY DETECTION IN DIFFERENT SIZE FRACTIONS

Because chromosome Y sequences appeared to be shorter (Fig. 2), we investigated whether selecting reads that had shorter lengths would increase the fetal DNA fraction and improve aneuploidy detection. We di-



**Fig. 1.** Effect of library preparation and sequencing on the length distribution of DNA.

$\lambda$  DNA was digested with *AfuI* and then paired-end sequenced. The number of sequenced fragments is plotted against length. Each black dot represents the mean number of reads in every 20-bp bin. The red line is a locally weighted logistic regression fit.

vided the reads into 3 groups by size: 30–150 bp, 150–170 bp, and 170–600 bp. Each group represented approximately one-third of the total paired reads.

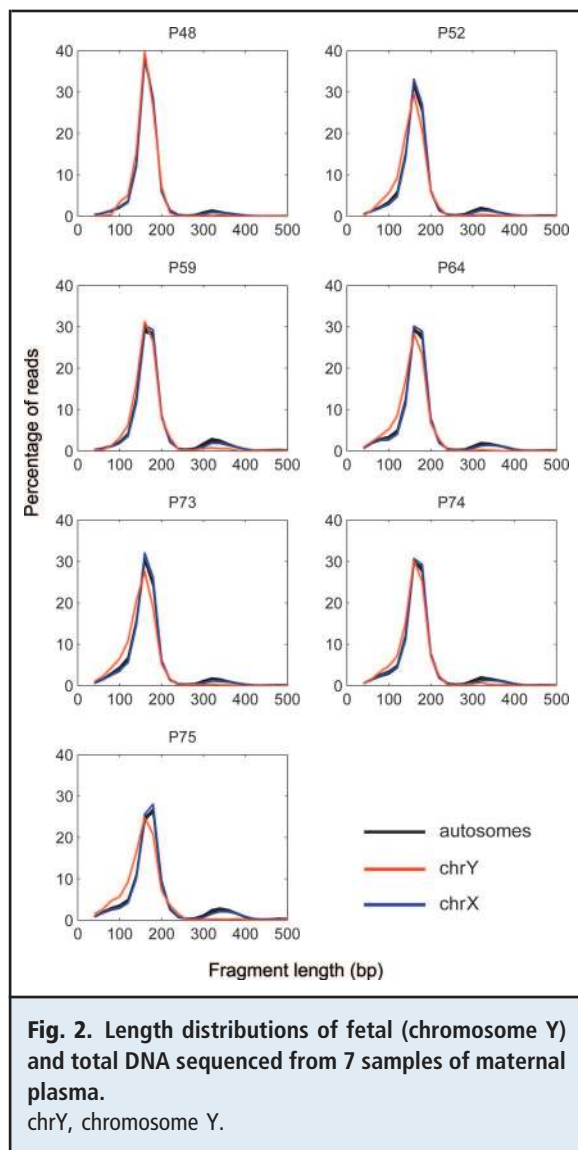
The fetal DNA percentage was calculated for all samples from the underrepresentation of chromosome X and/or the overrepresentation of trisomic chromosomes for all reads and for each size fraction, after correction for GC bias (Table 1). The fetal DNA percentage for the fraction of <150 bp was generally higher (by a factor of approximately 1.2–2) than the overall fetal DNA percentage (when all reads were taken into account), whereas for the fractions >150 bp, the fetal DNA percentage was lower than the overall value (Fig. 3). Thus, selecting reads of <150 bp was able to enrich the fetal DNA fraction.

We calculated the *z* statistic, a measure that reflects the confidence in the deviation of the representation of a chromosome from normal. Because the statistic also depends on the number of reads being considered, we randomly selected a third of the total reads within a sample for comparison. This random selection of reads had fragment sizes that represented the overall length distribution in the cell-free DNA sample. Although the fetal fraction and relative chromosome copy number were highest for the fraction of <150 bp, as observed by

the increase in the deviation of the relative copy number of chromosome X and trisomic chromosomes from 1.0 (Fig. 4A), the magnitude of the *z* statistic was not always the highest. In 4 of the 7 cases, the sensitivity was highest when all fractions were used (Fig. 4B).

## Discussion

We have demonstrated the direct measurement of the size distributions of maternal and fetal DNA in maternal plasma. A few recent studies also used traditional sequencing and 454 pyrosequencing to study the size distributions and profiles of cell-free DNA in healthy individuals and cancer patients (16–18). We also attempted to measure the size distribution of maternal cell-free DNA with 454 pyrosequencing (3). The results of these studies were in agreement that cell-free DNA has a peak size of 160–180 bp and that this DNA derives mostly from apoptotic cells. In this study, we used paired-end sequencing on the Illumina/Solexa platform, which has a much higher throughput than the 454 platform with respect to the number of reads sequenced. The large number of reads enabled us to characterize the size distribution of not only maternal DNA but also chromosome Y sequences, which consti-



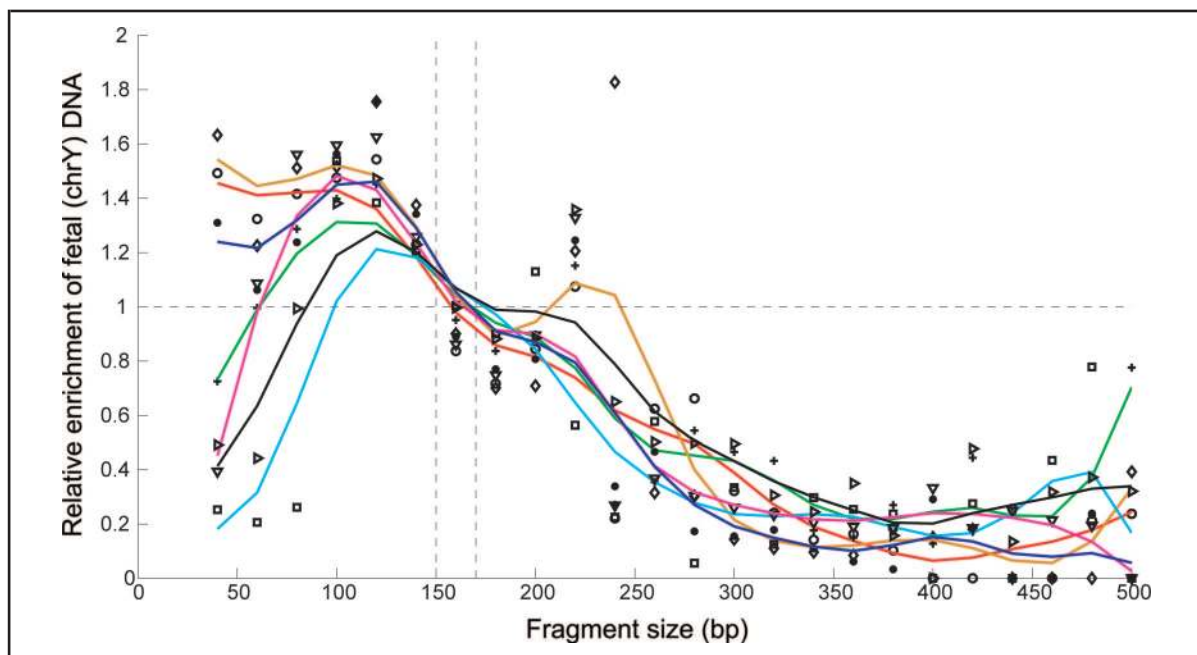
tuted only approximately 0.1% of all the sequences in a maternal plasma DNA sample. Our sequencing approach, however, could measure only the distribution in the lower molecular weight range because higher molecular weight species (>1 kb) undergo attrition in the current preparation of sequencing samples. Because previous experiments with the PCR and gel electrophoresis have shown that the majority of fetal DNA had sizes <500 bp, the current measurement approach should capture the size distribution of most fetal DNA. We noted that Southern blots of maternal plasma DNA revealed the presence of DNA with sizes >20 kb (7). Future experiments with the newly developed mate pair sample-preparation technique, which allows an insert size of >2 kb (19), should give a detailed size estimate of the higher molecular weight species.

In our previous study, the estimates of the fraction of fetal DNA obtained from sequencing data ranged from 8% to 40%, higher than the estimates from our own digital PCR measurements before sequencing library preparation and the estimates of <10% observed by others (20). Our explanations at that time had 2 components: (a) According to the studies of Li et al. (7) and Chan et al. (6), fetal DNA in maternal plasma is shorter than the maternal DNA counterpart; and (b) our sequencing method involved PCR amplifications with universal primers during library preparation and cluster generation. The PCR is known to have a higher efficiency for lower molecular weight species. We speculated that the increased fraction of fetal DNA measured from the sequencing data was an artifact of the sequencing method we chose to use.

In this study, we experimentally validated both components of our argument. By sequencing restriction digests of  $\lambda$  DNA, we discovered that lower molecular weight species were overrepresented. In addition, our sequencing measurements of the size distributions of maternal and fetal DNA (Fig. 2) agreed with previous findings that fetal DNA is mostly shorter than 300 bp, whereas a portion of maternal DNA is >300 bp in size (6, 7). These observations suggested that the process of sequencing maternal plasma DNA with the Solexa/Illumina platform increased the representation of the shorter fetal DNA species, thereby increasing the fetal DNA fraction.

Since the discovery that fetal DNA is generally shorter than maternal DNA in maternal plasma, a number of techniques have been developed to enrich fetal DNA fraction by size selection. These techniques have included traditional gel electrophoresis (7, 21), combinations of PCR assays with amplicons of different lengths (11), and microchip separation (22). Because the length bias of the shotgun sequencing reads was suspected to derive from the PCR, one could potentially ligate universal adaptors to the ends of plasma DNA and then perform PCR amplification against the universal sequences to enrich the fetal DNA fraction. In using these 2 approaches, it is important that the plasma DNA not be fragmented by nebulization or sonication so that the original size distribution of the DNA can be preserved.

The sensitivity of fetal aneuploidy detection via the counting of single DNA molecules depends on both the fetal DNA fraction and the number of molecules counted. Aneuploidy is more confidently detected if the fetal DNA fraction is high and a large number of molecules are counted. Our results show that although the fetal DNA fraction is increased in the shortest fragments (Fig. 4A), the fact that the total number of molecules being counted is smaller negatively affects the confidence of detection (Fig. 4B). Therefore, “infor-



**Fig. 3.** Relative increase or decrease in the fetal DNA fraction from 30 bp to 510 bp at 20-bp intervals.

We used locally weighted logistic regression to visualize the trend (solid line). Each patient sample is represented by a different symbol and a differently colored solid line. The vertical dashed lines represent the size cutoffs used to divide the reads into 3 portions with approximately equal numbers of reads. chrY, chromosome Y.

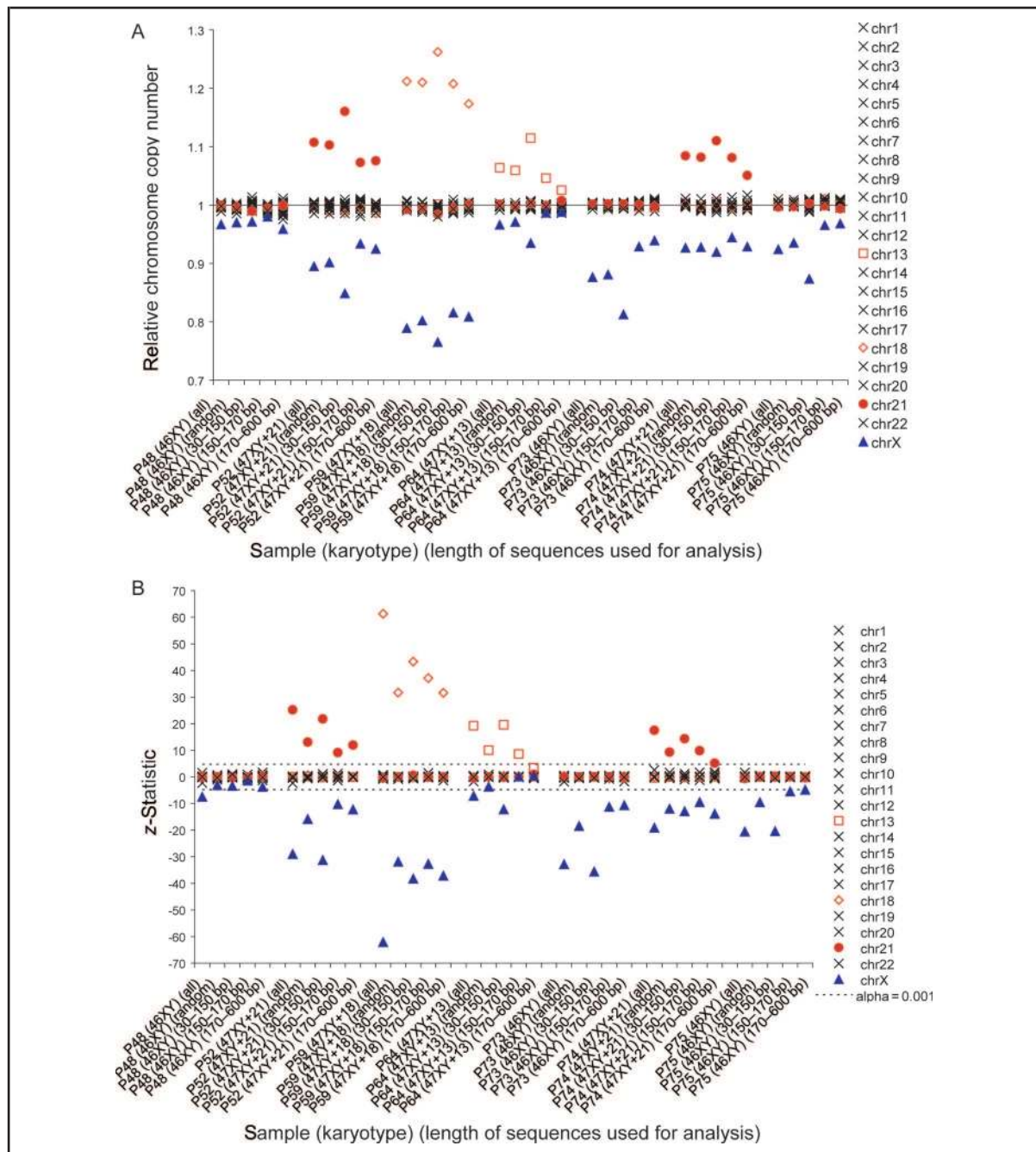
matic” enrichment of length fragments by digital size selection, whether by paired-end sequencing or by digital PCR (11, 23), may not yield any appreciable gain in the sensitivity of aneuploidy detection and should be used with caution. Whether one can gain sensitivity in aneuploidy detection depends on the initial fetal DNA fraction, the magnitude of the increase in the fetal DNA fraction obtained by size selection, and the number of

molecules retained after size selection. One situation in which we can imagine digital size selection being quite useful is when samples have been obtained suboptimally. For instance, if the processing of blood samples is delayed, maternal lymphocytes will lyse and artificially decrease the fetal fraction by contaminating the sample with longer fragments of maternal genomic DNA. These longer fragments potentially can be ex-

**Table 1.** Size distribution of fetal and total DNA in maternal plasma and fetal DNA percentages in different size fractions.

Sample	Karyotype	Median length, bp		Fetal DNA from chrX, %			Fetal DNA from trisomic chr, %				
		Non-chrY <sup>a</sup>	ChrY	All reads	0–150 bp	150–170 bp	170–600 bp	All reads	0–150 bp	150–170 bp	170–600 bp
P48	46XY	164	163	6.64	5.72	4.00	8.28	—	—	—	—
P52	47XY+21	161	153	21.06	30.30	13.32	15.04	21.46	32.08	14.58	15.16
P59	47XY+18	165	162	42.20	46.90	36.88	38.34	42.34	52.48	41.54	34.72
P64	47XY+13	164	155	6.78	13.08	2.68	2.76	12.72	22.88	9.22	5.12
P73	46XY	159	148	24.72	37.46	14.38	12.20	—	—	—	—
P74	47XY+21	164	159	16.44	16.52	11.48	14.32	16.82	22.18	16.20	10.16
P75	46XY	164	152	15.16	25.32	6.92	6.44	—	—	—	—

<sup>a</sup> chrY, chromosome Y.



**Fig. 4.** Copy number for each chromosome relative to the diploid condition and confidence in the measurement of chromosome representation in the different size fractions.

(A), Copy number of each chromosome relative to the diploid condition in different size fractions. Chromosome X (chrX) is underrepresented for all 7 male pregnancies. Chromosomes 13, 18, and 21 are overrepresented in the corresponding cases of trisomy 13, 18, and 21. The degree of deviation from the typical diploid representation is generally larger when analyzing reads <150 bp, compared with randomly selecting the similar number of reads from all reads, whereas the deviation decreases as the length of the analyzed reads increases. (B), Confidence in the measurement of chromosome representation in different size fractions. The horizontal dashed lines represents the 99.99% confidence level (2-tailed) for the detection of over- or underrepresentation of chromosomes. In 4 of the 7 cases, the sensitivity is highest when all fractions are used for analysis, although the fetal DNA percentage is generally higher in the size fraction of <150 bp.

cluded without reducing the number of fetal fragments used in the analysis.

In conclusion, we have shown that paired-end sequencing allows the direct measurement of the length distribution of cell-free fetal DNA in maternal plasma, with single-base resolution. The process of Illumina/Solexa sequencing introduced a bias in the length distribution of the sequenced sample that increased the representation of fetal DNA. Selecting sequenced reads with lengths <150 bp could further increase the fetal DNA fraction but would not necessarily increase the sensitivity of aneuploidy detection by single-molecule counting. We envision that the rapid advances in sequencing and related technologies will enable the realization of many novel techniques for the study of cell-free nucleic acids, not only for prenatal diagnosis but also for early cancer diagnosis.

**Author Contributions:** All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design,

acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

**Authors' Disclosures of Potential Conflicts of Interest:** Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

**Employment or Leadership:** None declared.

**Consultant or Advisory Role:** S.R. Quake, Fluidigm Corporation, Helicos Biosciences Corporation, and Artemis.

**Stock Ownership:** S.R. Quake, Fluidigm Corporation, Helicos Biosciences Corporation, and Artemis.

**Honoraria:** None declared.

**Research Funding:** The work was supported by the Wallace H. Coulter Foundation and the National Institutes of Health Director's Pioneer Award. H.C. Fan was supported by a Stanford graduate fellowship and an award from the Siebel Scholars Foundation.

**Expert Testimony:** None declared.

**Role of Sponsor:** The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

**Acknowledgments:** We thank the Division of Perinatal Genetics and General Clinical Research Center of Stanford University for patient recruitment and enrollment. We also thank Norma Neff for her help in performing sequencing experiments.

## References

1. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin No. 88, December 2007. Invasive prenatal testing for aneuploidy. *Obstet Gynecol* 2007;110:1459–67.
2. Dennis Lo YM, Chiu RW. Prenatal diagnosis: progress through plasma nucleic acids. *Nat Rev Genet* 2007;8:71–7.
3. Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc Natl Acad Sci U S A* 2008;105:16266–71.
4. Chiu RW, Chan KC, Gao Y, Lau VY, Zheng W, Leung TY, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci U S A* 2008;105:20458–63.
5. Chiu RW, Sun H, Akolekar R, Clouser C, Lee C, McKernan K, et al. Maternal plasma DNA analysis with massively parallel sequencing by ligation for noninvasive prenatal diagnosis of trisomy 21. *Clin Chem* 2010;56:459–63.
6. Chan KC, Zhang J, Hui AB, Wong N, Lau TK, Leung TN, et al. Size distributions of maternal and fetal DNA in maternal plasma. *Clin Chem* 2004;50:88–92.
7. Li Y, Zimmermann B, Rusterholz C, Kang A, Holzgreve W, Hahn S. Size separation of circulatory DNA in maternal plasma permits ready detection of fetal DNA polymorphisms. *Clin Chem* 2004;50:1002–11.
8. Li Y, Di Naro E, Vitucci A, Grill S, Zhong XY, Holzgreve W, Hahn S. Size fractionation of cell-free DNA in maternal plasma improves the detection of a paternally inherited beta-thalassemia point mutation by MALDI-TOF mass spectrometry. *Fetal Diagn Ther* 2009;25:246–9.
9. Li Y, Holzgreve W, Page-Christiaens GC, Gille JJ, Hahn S. Improved prenatal detection of a fetal point mutation for achondroplasia by the use of size-fractionated circulatory DNA in maternal plasma—case report. *Prenat Diagn* 2004;24:896–8.
10. Li Y, Wenzel F, Holzgreve W, Hahn S. Genotyping fetal paternally inherited SNPs by MALDI-TOF MS using cell-free fetal DNA in maternal plasma: influence of size fractionation. *Electrophoresis* 2006;27:3889–96.
11. Lun FM, Tsui NB, Chan KC, Leung TY, Lau TK, Charoenkwan P, et al. Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. *Proc Natl Acad Sci U S A* 2008;105:19920–5.
12. Sikora A, Zimmermann BG, Rusterholz C, Birri D, Kolla V, Lapaire O, et al. Detection of increased amounts of cell-free fetal DNA with short PCR amplicons. *Clin Chem* 2009;56:136–8.
13. Kozarewa I, Ning Z, Quail MA, Sanders MJ, Beriman M, Turner DJ. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat Methods* 2009;6:291–5.
14. White RA 3rd, Blainey PC, Fan HC, Quake SR. Digital PCR provides sensitive and absolute calibration for high throughput sequencing. *BMC Genomics* 2009;10:116.
15. Fan HC, Quake SR. Sensitivity of noninvasive prenatal detection of fetal aneuploidy from maternal plasma using shotgun sequencing is limited only by counting statistics. *PLoS ONE* 2010;5:e10439.
16. van der Vaart M, Semenov DV, Kuligina EV, Richter VA, Pretorius PJ. Characterisation of circulating DNA by parallel tagged sequencing on the 454 platform. *Clin Chim Acta* 2009;409:21–7.
17. Suzuki N, Kamataki A, Yamaki J, Homma Y. Characterization of circulating DNA in healthy human plasma. *Clin Chim Acta* 2008;387:55–8.
18. Beck J, Urnovitz HB, Riggert J, Clerici M, Schutz E. Profile of the circulating DNA in apparently healthy individuals. *Clin Chem* 2009;55:730–8.
19. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 2008;456:53–9.
20. Lo YM, Tein MS, Lau TK, Haines CJ, Leung TN, Poon PM, et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am J Hum Genet* 1998;62:768–75.
21. Jorgez CJ, Bischoff FZ. Improving enrichment of circulating fetal DNA for genetic testing: size fractionation followed by whole gene amplification. *Fetal Diagn Ther* 2009;25:314–9.
22. Hahn T, Drese KS, O'Sullivan CK. Microsystem for isolation of fetal DNA from maternal plasma by preparative size separation. *Clin Chem* 2009;55:2144–52.
23. Fan HC, Quake SR, inventors. Method of characterizing single DNA molecules based on length. US provisional patent application 12/541,271.